

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-117584

(43)Date of publication of application : 27.04.2001

(51)Int.Cl.

G10L 15/18
 G01C 21/00
 G08G 1/0969
 G10L 15/00
 G10L 15/28
 G10L 15/02
 // G10L101:00
 G10L101:065

(21)Application number : 11-300809

(71)Applicant : ALPINE ELECTRONICS INC

(22)Date of filing : 22.10.1999

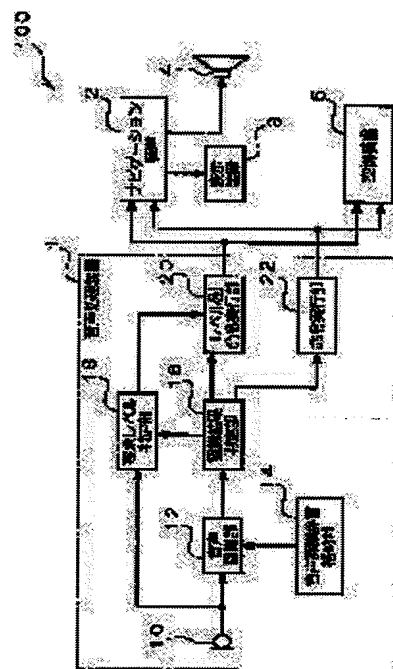
(72)Inventor : OISHI TETSUYA
 SAITO FUMIO

(54) VOICE PROCESSOR

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a voice processor which can make a proper operation instruction even when a word whose operation object and request level are not shown concretely is inputted.

SOLUTION: A recognition result judging part 16 judges whether or not a specified character string after the voice recognition by a voice recognition part 12 includes a desire type character string (operation character string of unconcrete indication type) whose operation object is not specified. A request level judging part 18 detects variation in sound pressure level according to a sound signal to judge the request level regarding the desire type character string. When a recognition result judging part 16 makes an affirmative judgment, an execution-by-level part 20 specifies an operation object according to the contents of the desiring type character string and the request level judging result and gives a specific operation instruction.



LEGAL STATUS

[Date of request for examination]

29.07.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-117584

(P2001-117584A)

(43) 公開日 平成13年4月27日 (2001.4.27)

(51) Int.Cl.⁷

識別記号

F I

テーマコード* (参考)

G 1 0 L 15/18

G 0 1 C 21/00

H 2 F 0 2 9

G 0 1 C 21/00

G 0 8 G 1/0969

5 D 0 1 5

G 0 8 G 1/0969

G 1 0 L 101:00

5 H 1 8 0

G 1 0 L 15/00

101:065

9 A 0 0 1

15/28

3/00

5 3 7 Z

審査請求 未請求 請求項の数 5 O L (全 10 頁) 最終頁に続く

(21) 出願番号

特願平11-300809

(22) 出願日

平成11年10月22日 (1999. 10. 22)

(71) 出願人 000101732

アルパイン株式会社

東京都品川区西五反田1丁目1番8号

(72) 発明者 大石 哲也

東京都品川区西五反田1丁目1番8号 ア
ルパイン株式会社内

(72) 発明者 斉藤 文男

東京都品川区西五反田1丁目1番8号 ア
ルパイン株式会社内

(74) 代理人 100103171

弁理士 雨貝 正彦

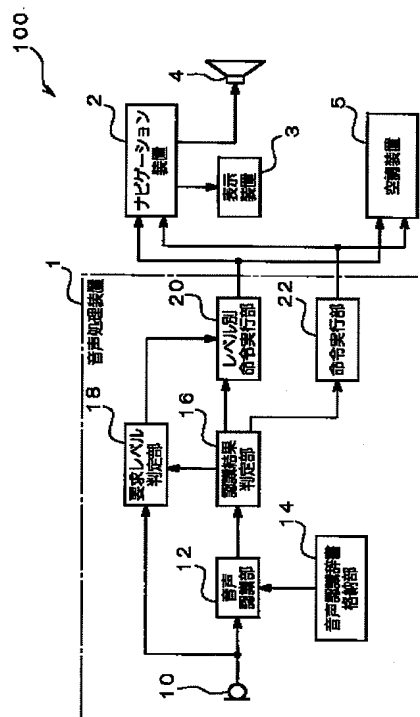
最終頁に続く

(54) 【発明の名称】 音声処理装置

(57) 【要約】

【課題】 操作対象と要求レベルが具体的に示されていない言葉が入力された場合でも適切な動作指示を行うことができる音声処理装置を提供すること。

【解決手段】 認識結果判定部16は、音声認識部12により音声認識処理が行われて特定された文字列に操作対象を特定しない欲求型文字列（非具体指示型の操作文字列）が含まれているか否かを判定する。また、要求レベル判定部18は、音声信号に基づいて音圧レベルの変化を検出して欲求型文字列に関する要求レベルを判定する。認識結果判定部16により肯定的な判断がなされた場合に、レベル別命令実行部20は、欲求型文字列の内容と要求レベル判定結果とに基づいて操作対象を特定して所定の動作指示を行う。



【特許請求の範囲】

【請求項 1】 入力された音声信号に対して音声認識処理を行って、前記音声信号に対応する認識文字列を特定する音声認識手段と、

前記音声認識手段によって特定された前記認識文字列が、操作対象を特定しない非具体指示型の操作文字列を含むか否かを判定する判定手段と、

前記音声信号および前記認識文字列の少なくとも一方に基づいて前記操作文字列に関する要求レベルを判定するレベル判定手段と、

前記判定手段によって肯定的な判断がなされたときに、前記操作文字列の内容と前記要求レベルとに基づいて操作対象を特定して、所定の動作指示を行う動作指示手段と、

を備えることを特徴とする音声処理装置。

【請求項 2】 請求項 1 において、前記判定手段は、前記認識文字列に前記非具体指示型の操作文字列が含まれるか否かの判定とともに、操作対象を特定した具体指示型の操作文字列が含まれるか否かの判定を行っており、

前記動作指示手段は、前記判定手段によって前記認識文字列に前記具体指示型の操作文字列が含まれると判断されたときに、この操作文字列によって特定される操作対象に対して所定の動作指示を行うことを特徴とする音声処理装置。

【請求項 3】 請求項 1 または 2 において、前記認識文字列に前記非具体指示型の操作文字列が含まれている場合に前記動作指示手段によって行われる動作指示候補には、情報の提示と動作の実行とが少なくとも含まれており、前記レベル判定手段によって判定された要求レベルの程度が低い場合に、前記動作指示手段は、特定された前記操作対象に対して前記情報の提示を指示することを特徴とする音声処理装置。

【請求項 4】 請求項 1～3 のいずれかにおいて、前記レベル判定手段は、前記音声信号に基づいて声の大きさ、声の高さおよび発声間隔の少なくとも 1 つを特徴量として抽出し、抽出した前記特徴量に基づいて前記要求レベルを判定することを特徴とする音声処理装置。

【請求項 5】 請求項 1～3 のいずれかにおいて、前記レベル判定手段は、前記認識文字列に前記要求レベルを表す所定の修飾語に関する文字列が含まれている場合に、この文字列に基づいて前記要求レベルを判定することを特徴とする音声処理装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 本発明は、入力される音声に対して音声認識を行い、音声認識結果に基づいて各種の動作指示命令を出力する音声処理装置に関する。

【0002】

【従来の技術】 従来から、音声認識技術を用いて各種装

置に対する操作指示等を入力する各種システムが知られている。このようなシステムでは、操作指示等を入力するために利用者が発声した音声に対して音声認識処理を行って操作対象の装置と操作内容を特定し、対応する動作指示命令を出力するための音声処理装置が備えられている。例えば、ナビゲーション装置、オーディオ装置、空調装置等の各種車載用装置により構成される車載用システムを例にとると、上述したような音声処理装置を備えることにより、各装置に対して音声で操作指示を与えることができるので、特に走行中に体が振動するような場合における操作が容易となる。

【0003】 上述した従来の音声処理装置では、音声による操作指示を与える際には、操作対象の装置と操作内容を明確に示してやる必要がある。例えば、上述した車載用システムの例において自車位置周辺にあるレストランを探したい場合であれば、利用者が「レストランを探す」等の音声を入力することにより、音声処理装置は、「レストラン」から操作指示の対象がナビゲーション装置であり、「探す」から具体的な指示内容が自車位置周辺にあるレストランの検索であることを特定し、その旨を示す動作指示命令をナビゲーション装置に対して出力する。

【0004】

【発明が解決しようとする課題】 ところで、一般的に、何らかの心情や欲求等を感じた人間が最初に発する言葉は、その心情や欲求等を素直に表現したものである場合が多い。これは、まだ語彙の少ない子供が発する言葉を見ればよく理解できるように、例えば、空腹を感じた場合であれば、「〇〇が食べたい」、「レストランに行きたい」等の具体的な内容を含んだ言葉よりも先に、まず、「お腹がすいた」、「腹へった」等の欲求が率直に表現された言葉が発声されるのが普通である。ところが、従来の音声処理装置では、このように心情や欲求等が率直に表現された言葉からは操作対象の装置と操作内容を特定することができないので、利用者は、操作指示を入力するために必要な特定の言葉をあらかじめ覚えておく必要があった。このため、従来の音声処理装置は、必ずしも利用者に使いやすいものとはいえなかった。従来の音声処理装置において上述したような心情や欲求等が率直に表現された言葉に対応するには、これらの言葉と具体的な動作指示とをあらかじめ対応づけて記憶しておく必要がある。例えば、「お腹がすいた」という言葉に対して「レストランの検索」という動作指示、「暑い」という言葉に対して「空調装置の温度設定を下げる」という動作指示、「疲れた」という言葉に対して「休憩地の検索」という動作指示、といった形で言葉と動作指示を一对一に対応づけておけば、心情や欲求等が率直に表現された言葉に対しても従来の音声処理装置で対応することができる。しかしながら、例えば「お腹がすいた」という言葉を例にとって考えてみても、この言

業を発した利用者が有している要求レベルは、例えば、今すぐに最寄りのコンビニエンスストア等を探して何かを買って食べたいという場合もあるし、少々はがまんできるので美味しいレストランを探して食事したいという場合も考えられる。ところが、上述したように、言葉と動作指示とが一对一に対応づけられていると、「お腹がすいた」という言葉に対して行われる動作指示は常に「レストランの検索」となり、「最寄りのコンビニエンスストアの検索」という動作指示が行われることはない。すなわち、従来の音声処理装置では、発せられた言葉から利用者の要求レベルを判定して、その要求レベルに対応した動作指示を行うことはできなかった。

【0005】本発明は、このような点に鑑みて創作されたものであり、その目的は、操作対象と要求レベルが具体的に示されていない言葉が入力された場合でも適切な動作指示を行うことができる音声処理装置を提供することにある。

【0006】

【課題を解決するための手段】上述した課題を解決するために、本発明の音声処理装置は、音声信号が入力された際に、音声認識手段により音声認識処理を行って音声信号に対応する認識文字列を特定し、判定手段によってこの認識文字列に操作対象を特定しない非具体指示型の操作文字列が含まれているか否かを判定するとともに、レベル判定手段により音声信号および認識文字列の少なくとも一方に基づいて操作文字列に関する要求レベルを判定しており、判定手段により肯定的な判断がなされた場合には、操作文字列の内容とレベル判定手段によって判定された要求レベルとに基づき動作指示手段により操作対象を特定して、所定の動作指示を行っている。したがって、操作対象と要求レベルが具体的に示されていないような音声が入力された場合でも、操作文字列の内容とこの操作文字列に関する要求レベルとに基づいて操作対象を特定して動作指示を行うことができるので、人間同士の会話に近い形の自然な言葉を用いて、各種の操作指示を入力することができるとともに、要求レベルに応じた適切な動作指示を選択することができる。

【0007】また、上述した判定手段は、認識文字列に非具体指示型の操作文字列が含まれるか否かの判定とともに、操作対象を特定した具体指示型の操作文字列が含まれるか否かの判定を行い、認識文字列に具体指示型の操作文字列が含まれている場合には、動作指示手段によって、具体指示型の操作文字列に基づいて特定される操作対象に対して所定の動作指示を行うことが望ましい。認識文字列に具体指示型の操作文字列が含まれている場合には、この具体指示型の操作文字列から操作対象を簡単に特定することができるので、動作指示を行うための処理を簡略化することができる。

【0008】また、認識文字列に非具体指示型の操作文字列が含まれている場合において、レベル判定手段によ

って判定された要求レベルの程度が低い場合に、動作指示手段は、非具体指示型の操作文字列により特定された操作対象に対して各種の情報の提示を行うよう指示することが望ましい。要求レベルの程度が低い場合には、操作対象に対して積極的に動作を要求しているわけではなく各種の情報の提示のみを要求している場合もあるが、このような場合にも対応することができ、利用者の要求に対してより細かく対応した動作指示を行うことができる。

【0009】また、上述したレベル判定手段は、音声信号に基づいて要求レベルを判定する場合には、声の大きさ、声の高さおよび発声間隔の少なくとも1つを特徴量として抽出し、この特徴量に基づいて要求レベルを判定することが望ましい。一般的に、感情の変化は音声の変化として現れやすいので、音声信号から上述した各特徴量の少なくとも1つを抽出することにより、利用者の有する要求レベルを容易に判定することができる。

【0010】また、上述したレベル判定手段は、認識文字列に基づいて要求レベルを判定する場合には、認識文字列に含まれる所定の修飾語に関する文字列に基づいて要求レベルを判定することが望ましい。例えば、認識文字列に「少し」、「とても」等の修飾語に関する文字列が含まれている場合には、これらの文字列に基づいて要求レベルを判定することができる。したがって、音声の変化が少なく音声信号から抽出した各特徴量の変化による要求レベルの判定が困難な利用者に対しても、上述したような所定の修飾語を判断要素とすることにより利用者の有する要求レベルを確実に判定することができる。

【0011】

【発明の実施の形態】以下、本発明の音声処理装置を車載用システムに適用した一実施形態について、図面を参照しながら説明する。図1は、一実施形態の車載用システムの構成を示す図である。図1に示す車載用システム100は、利用者が発声した音声に対応して各種の動作指示を行う音声処理装置1と、自車位置を検出して自車位置周辺の地図を表示したり、利用者によって選択された目的地までの経路探索および経路誘導等を行うナビゲーション装置2と、ナビゲーション装置2から出力される各種画像を表示する表示装置3と、ナビゲーション装置2から出力される案内音声等の各種音声出力するスピーカ4と、車室内の温度等を調節する空調装置5を含んで構成されている。

【0012】図1に示す音声処理装置1は、マイクロホン10、音声認識部12、音声認識辞書格納部14、認識結果判定部16、要求レベル判定部18、レベル別命令実行部20、命令実行部22を含んで構成されている。

【0013】マイクロホン10は、利用者が発声した音声を集音して音声信号に変換する。音声認識部12は、マイクロホン10から出力される音声信号を解析して所

定の音声認識処理を行い、利用者が発声した音声に対応する文字列を特定する。音声認識辞書格納部14は、標準的な音声に対応した信号波形を音声認識用の辞書として格納している。

【0014】認識結果判定部16は、音声認識部12によって特定された文字列（認識文字列）が、利用者の欲求を表現したものであるか、あるいは動作指示命令を与える対象となる装置と行わせたい動作内容を明確に示したものであるかを判定する。以後、上述した利用者の欲求を表現した文字列を「欲求型文字列」と称し、動作指示を行う対象となる装置と行わせたい動作内容を明確に示した文字列を「具体指示型文字列」と称するものとする。なお、この「欲求型文字列」が「非具体指示型の操作文字列」に対応しており、「具体指示型文字列」が「具体指示型の操作文字列」に対応している。認識結果判定部16は、文字列が欲求型文字列であると判定した場合には、この文字列をレベル別命令実行部20に出力するとともに、要求レベル判定部18に対して要求レベル検出結果を出力するよう指示を行う。

【0015】図2は、欲求型文字列と具体指示型文字列の一例を示す図である。図2では、カタカナで表現されているものが音声認識結果としての文字列を示しており、このカタカナで表現された文字列の下段にカッコ書きで表現されているのが利用者が実際に発声した音声を示している。図2に示すように、欲求型文字列としては、「ツカレタ（疲れた）」、「オナカガスイタ（お腹がすいた）」、「アツイ（暑い）」、「サムイ（寒い）」などが挙げられる。

【0016】また、具体指示型文字列としては、「エアコンヲツケル（エアコンをつける）」、「レストランヲサガス（レストランを探す）」、「ヌケミチヲサガス（抜け道を探す）」などが挙げられる。具体的には、例えば、上述した具体指示型文字列の1つである「エアコンヲツケル（エアコンをつける）」では、空調装置を示す文字列「エアコン」から動作指示命令を与える対象が空調装置5であることを特定することができ、文字列「ツケル」から動作内容（空調装置5を動作させる）を特定することができる。他にも、具体指示型文字列の1つである「ヌケミチヲサガス（抜け道を探す）」では、文字列「ヌケミチ」および文字列「サガス」の各々から、ナビゲーション装置2（対象装置）に抜け道探索（動作内容）を行わせたいということを特定することができる。このように、具体指示型文字列は、動作指示を行う対象となる装置と行わせたい動作内容が明確に示されているという特徴を有する。なお、欲求型文字列および具体指示型文字列は、上述した一例に限定されるものではなく、他にも様々なパターンが考えられる。

【0017】ところで、実際の音声認識処理では「エアコンをつける」等の入力音声に対して抽出される文字列は、「を」等の助詞を省略して「エアコン」と「ツケ

ル」の2つの文字列として特定されるのが一般的であるが、本明細書では説明を簡略化するために、音声認識処理によって特定される文字列の中に「を」等の助詞に対応する文字列も含めてひとまとまりの文字列として表現するものとする。

【0018】要求レベル判定部18は、利用者が発声した音声の音圧レベルの変化（声の大きさの変化）に基づいて利用者の要求レベルを判定する。図3は、要求レベル判定部18の詳細構成を示す図である。図3に示すように、要求レベル判定部18は、音圧レベル測定部180、平均音圧レベル記憶部182、音圧レベル比較部184を含んで構成されている。音圧レベル測定部180は、マイクロホン10から出力される音声信号に基づいて音圧レベルを測定し、測定された音圧レベルをデジタルデータに変換する。平均音圧レベル記憶部182は、利用者により発声される音声の平均的な音圧レベルをデジタルデータとして記憶している。この平均的な音圧レベルは、利用者が本実施形態の音声処理装置1に対して音声を入力する毎に更新される。音圧レベル比較部184は、音圧レベル測定部180から出力される音圧レベルを平均音圧レベル記憶部182に記憶されている平均的な音圧レベルと比較することにより、利用者の有している要求レベルをレベル1～3の3段階に分類する。具体的には、音圧レベル比較部184は、音圧レベル測定部180から出力された音圧レベルが平均的な音圧レベルよりも低い場合をレベル1、平均的な音圧レベルと同程度の場合をレベル2、平均的な音圧レベルよりも高い場合をレベル3と分類する。このレベル値が利用者の要求レベルの高さに対応している。すなわち、レベル1は要求レベルが低いことを表し、レベル3は要求レベルが高いことを表し、レベル2は要求レベルがレベル1とレベル3の間であることを表している。

【0019】レベル別命令実行部20は、ナビゲーション装置2等を動作させるための各種の動作指示が選択候補として動作指示テーブルに格納されており、認識結果判定部16から出力される文字列（欲求型文字列）と要求レベル判定部18から出力される要求レベル判定結果に対応して適切な動作指示を選択して実行する。図4は、動作指示テーブルの一例を示す図であり、欲求型文字列の内容に対応してレベル1～3にレベル分けされて用意された動作指示が格納されている。例えば、「ツカレタ」という欲求型文字列に対して、レベル1：「連続運転時間の提示」、レベル2：「休憩を促すメッセージの提示」、レベル3「休憩地の探索指示」というようにレベル分けされた動作指示が格納されている。また、「アツイ」という欲求型文字列に対しては、レベル1：「車室内温度の提示」、レベル2：「空調装置を動作させるかを問うメッセージの提示」、レベル3：「空調装置の動作指示」というような動作指示が格納されている。なお、欲求型文字列と対応する動作指示は図4に示

した一例に限定されるものではなく、他にも様々なパターンが考えられる。

【0020】レベル別命令実行部20は、認識結果判定部16から出力された欲求型文字列が「ツカレタ」であり、要求レベル判定部18から出力された利用者の要求レベルがレベル2である場合には、上述した図4に示した動作指示テーブルの中から「休憩を促すメッセージの提示」を選択して実行する。この場合には、利用者に対して「休憩をとられてはいかがですか？」等のメッセージが、画像および音声により提示される。このような画像および音声によるメッセージの提示を行う場合に、レベル別命令実行部20は、ナビゲーション装置2に備えられた画像生成部（図示せず）やオーディオ部（図示せず）等を利用して画像および音声信号を生成し、表示装置3およびスピーカ4を用いてこれらの画像を表示し、音声を出力する。

【0021】命令実行部22は、認識結果判定部16から出力される具体指示型文字列に対応した各種の動作指示を格納しており、具体指示文字列の内容に応じた動作指示命令を選択してナビゲーション装置2等に向けて出力する。例えば、上述した図2に示した具体指示型文字列の1つである「レストランヲサガス（レストランを探す）」が入力された場合であれば、命令実行部22は、ナビゲーション装置2に向けて、例えば、自車位置から半径5km以内にあるレストランを検索して表示するよう指示する動作指示を出力する。

【0022】上述した音声認識部12、音声認識用辞書格納部14が音声認識手段に、認識結果判定部16が判定手段に、要求レベル判定部18がレベル判定手段に、レベル別命令実行部20、命令実行部22が動作指示手段にそれぞれ対応している。

【0023】本実施形態の車載用システム100および音声処理装置1は上述した構成を有しており、次に、音声処理装置1が利用者から発声された音声に対して所定の処理を行ってナビゲーション装置2等に向けて動作指示を出力する際の動作手順について説明する。例えば、利用者が「疲れた」と発声した場合における音声処理装置1の動作手順について説明する。図5は、音声処理装置1の動作手順を示す流れ図である。

【0024】音声認識部12は、マイクロホン10に対して利用者の音声が入力された否かを常に判定しており（ステップ100）、音声入力となされた場合には、音声認識辞書格納部14に格納された音声認識用辞書を用いて、マイクロホン10によって集音された音声に対応する文字列を特定する（ステップ101）。上述したように利用者によって発声された音声「疲れた」であれば、音声認識部12はこの音声に対応する文字列「ツカレタ」を特定する。

【0025】音声認識部12によって文字列が特定されて出力されると、認識結果判定部16は、この文字列が

欲求型文字列であるか否かを判定する（ステップ102）。例えば、上述した文字列「ツカレタ」は、図2に示したテーブルを参照することにより欲求型文字列であると判定される。入力された文字列が欲求型文字列であると判定した場合に、認識結果判定部16は、この文字列をレベル別命令実行部20に出力する（ステップ103）とともに、要求レベル判定部18に対して、利用者の音声から検出した要求レベル判定結果をレベル別命令実行部20に出力するように指示する。

【0026】次に、レベル別命令実行部20は、認識結果判定部16から出力される文字列（欲求型文字列）と要求レベル判定部18から出力される要求レベル判定結果に対応して適切な動作指示を選択する（ステップ104）。例えば、認識結果判定部16から文字列「ツカレタ」が出力された場合において、要求レベル判定結果が「レベル1」である場合には、レベル別命令実行部20は、「連続運転時間の提示」の動作指令を選択する。また、レベル別命令実行部20は、要求レベル検出結果が「レベル2」であれば「休憩を促すメッセージの提示」の動作指示を選択し、「レベル3」であれば「休憩地の探索指示」の動作指示を選択する。

【0027】その後、レベル別命令実行部20は、選択した動作指示を実行する（ステップ105）。具体的には、上述した「連続運転時間の提示」（レベル1に対応）が選択された場合であれば、レベル別命令実行部20は、連続運転時間を画像および音声により利用者に対して提示するようナビゲーション装置2に対して指示を出力する。これにより、例えば「連続運転時間が2時間45分になりました」等のメッセージが表示装置3に表示されるとともにスピーカ4を介して音声出力され、利用者に対して提示される。同様に、「休憩を促すメッセージの提示」（レベル2に対応）が選択された場合であれば、「休憩をとられてはいかがですか？」等のメッセージが表示装置3およびスピーカ4を介して利用者に対して提示される。また、「休憩地の探索指示」（レベル3に対応）が選択された場合であれば、レベル別命令実行部20は、ナビゲーション装置2に対して、「休憩地を探索します」等のメッセージを表示装置3およびスピーカ4を介して利用者に対して提示するよう指示を行うとともに、自車位置周辺にある休憩地（例えば、コンビニエンスストアやパーキングエリア等）を検索するよう指示を行う。選択した動作指示に対応した動作が完了すると、ステップ100に戻り、音声入力となされたか否かの判定以降の動作が繰り返される。

【0028】また、音声認識部12から出力された文字列が欲求型文字列でないと判定された場合には、上述したステップ102において否定判断となされ、次に認識結果判定部16は、音声認識部12から出力された文字列が具体指示型文字列であるか否かを判定する（ステップ106）。文字列が具体指示型文字列でないと判定さ

れた場合には、入力された音声があらかじめ格納されている欲求型文字列および具体指示型文字列のいずれにも該当しないということなので、認識結果判定部16は、命令実行部22を介して画像および音声により利用者に対してエラー通知を行い、再度音声を入力するよう促す(ステップ107)。

【0029】また、音声認識部12から出力された文字列が具体指示型文字列である場合には、上述したステップ106において肯定判断がなされ、次に認識結果判定部16は、この文字列を命令実行部22に出力する(ステップ108)。命令実行部22は、認識結果判定部16から出力された文字列(具体指示文字列)の内容に応じた動作指示命令を選択する(ステップ109)。その後、命令実行部22は、選択された動作指示を実行する(ステップ110)。具体的には、例えば、認識結果判定部16から出力された文字列が「エアコンヲツケル」の場合には、命令実行部22は、空調装置5に対して動作指示を行い、空調装置5を作動させるとともに、動作内容を利用者に対して通知するために、「エアコンをおつけします」等のメッセージを表示装置3およびスピーカ4を介して提示する。選択された動作指示に対応した動作が完了すると、ステップ100に戻り、音声入力となされたか否かの判定以降の動作が繰り返される。

【0030】このように、本実施形態の音声処理装置1は、音声認識処理が行われて特定された文字列が欲求型文字列であるか否かを判定するとともに、音声信号から抽出される音圧レベルの変化に基づいて上述した欲求型文字列に関する要求レベルを判定し、欲求型文字列の内容と要求レベル判定結果とに基づいて操作対象を特定して所定の動作指示を行っている。したがって、上述した「疲れた」の例のように、操作対象と要求レベルが具体的に示されていないような音声が入力された場合においても、操作対象と要求レベルを特定して適切な動作指示を行うことができる。

【0031】なお、本発明は上記実施形態に限定されるものではなく、本発明の要旨の範囲で種々の変形実施が可能である。例えば、上述した実施形態では、要求レベル判定部18は、利用者が発声した音声の音圧レベルの変化に基づいて利用者の要求レベルを判定していたが、声の高さや発声間隔等の変化に基づいて利用者の要求レベルを判定するようにしてもよい。

【0032】図6は、声の高さの変化に基づいて利用者の要求レベルを判定する場合の要求レベル判定部18aの詳細構成を示す図である。図6に示すように、要求レベル判定部18aは、基本周波数測定部280、平均基本周波数記憶部282、基本周波数比較部284を含んで構成されている。基本周波数測定部280は、マイクロホン10から出力される音声信号に基づいて音声の基本周波数を測定し、測定された基本周波数をデジタルデータに変換する。この基本周波数は、一般的に、男性

の音声に対して100～200Hz程度、女性の音声に対して200～400Hz程度の値を有しており、この基本周波数の変化が声の高さの変化に対応する。すなわち、声が高い場合には基本周波数が高くなり、声が低い場合には基本周波数が低くなる。平均基本周波数記憶部282は、利用者により発声される音声の平均的な基本周波数をデジタルデータとして記憶している。この平均的な基本周波数は、利用者が音声を入力する毎に更新される。基本周波数比較部284は、基本周波数測定部280によって測定される基本周波数を平均基本周波数記憶部282に記憶されている平均的な基本周波数と比較することにより、利用者の有している要求レベルをレベル1～3の3段階に分類する。

【0033】具体的には、例えば、基本周波数比較部284は、基本周波数測定部280から出力された基本周波数が平均的な基本周波数よりも低い場合をレベル1、平均的な基本周波数と同程度の場合をレベル2、平均的な基本周波数よりも高い場合をレベル3と分類する。上述したように、このレベル値が利用者の要求レベルの高さに対応しており、レベル1は要求レベルが低いことを表し、レベル3は要求レベルが高いことを表し、レベル2は要求レベルがレベル1とレベル3の間であることを表している。なお、図6に示した要求レベル判定部18aでは、音声の基本周波数のみに着目して声の高さを検出していたが、音声の中の他の周波数成分も考慮して声の高さを検出するようにしてもよい。

【0034】図7は、発声間隔の変化に基づいて利用者の要求レベルを判定する場合の要求レベル判定部18bの詳細構成を示す図である。図7に示すように、要求レベル判定部18bは、発声間隔測定部380、平均発声間隔記憶部382、発声間隔比較部384を含んで構成されている。発声間隔測定部380は、マイクロホン10から出力される音声信号に基づいて音声の発声間隔を測定し、測定された発声間隔をデジタルデータに変換する。実際には、例えば「疲れた」という音声が入力された場合であれば、文字列「ツカレタ」に対して「ツ」、「カ」、「レ」、「タ」の各々の音声区間に対応する時間が、例えば数十ミリ秒程度の値として検出され、各音声区間に対応する時間の平均値として発声間隔が求められる。なお、発声間隔は、上述した例と異なる方法により求めてもよい。平均発声間隔記憶部382は、利用者により発声される音声の平均的な発声間隔をデジタルデータとして記憶している。ここでいう平均的な発声間隔とは、利用者が音声を複数回発声した場合の発声間隔の平均値を表している。この平均的な発声間隔は、利用者が音声を入力する毎に更新される。発声間隔比較部384は、発声間隔測定部380によって測定される発声間隔を平均発声間隔記憶部382に記憶されている平均的な発声間隔と比較することにより、利用者の有している要求レベルをレベル1～3の3段階に分類する。

【0035】具体的には、例えば、発声間隔比較部384は、発声間隔測定部380から出力された発声間隔が平均的な発声間隔よりも長い場合をレベル1、平均的な発声間隔と同程度の場合をレベル2、平均的な発声間隔よりも短い場合をレベル3と分類する。上述したように、このレベル値が利用者の要求レベルの高さに対応しており、レベル1は要求レベルが低いことを表し、レベル3は要求レベルが高いことを表し、レベル2は要求レベルがレベル1とレベル3の間であることを表している。

【0036】また、利用者が発声した音声の中に所定の修飾語が含まれているとき、この修飾語の内容に基づいて利用者の要求レベルを判定するようにしてもよい。図8は、この変形例における音声処理装置1aの構成を示す図である。図8に示す音声処理装置1aは、図1に示した音声処理装置1と比べて、要求レベル判定部18cが音声認識部12から出力される文字列に基づいて利用者の要求レベルを判定している点が異なっている。また、図9は、要求レベル判定部18cの詳細構成を示す図である。図9に示すように、要求レベル判定部18cは、修飾語検出部480、要求レベル認識用データ格納部482、要求レベル認識部484を含んで構成されている。

【0037】修飾語検出部480は、音声認識部12から出力される文字列の中に、所定の修飾語が含まれている場合にこれを検出する。所定の修飾語としては、例えば、「少し」、「ちょっと」、「少々」、「すごく」、「とても」、「かなり」等が考えられる。例えば、「すごく疲れた」という音声が入力され、この音声に対して音声認識部12により「スゴクツカレタ」という文字列が抽出された場合であれば、修飾語検出部480は、修飾語「すごく」に対応する文字列「スゴク」を検出する。また、所定の修飾語に対応する文字列が検出されなかった場合には、その旨を要求レベル認識部484に通知する。要求レベル認識用データ格納部482は、上述した所定の修飾語とこの修飾語により表される利用者の要求レベルとの関係を示すデータテーブルを格納している。

【0038】図10は、修飾語と利用者の要求レベルとの関係を示すデータテーブルの一例を示す図である。図10に示すように、上述した修飾語の「少し」、「ちょっと」、「少々」等は、利用者の要求レベルが低い場合に対応しており、レベル1に分類されている。また、上述した修飾語の「すごく」、「とても」、「かなり」等は、利用者の要求レベルが高い場合に対応しており、レベル3に分類されている。また、利用者が発声した音声に上述したような修飾語が付加されていない場合がレベル2に分類されている。

【0039】要求レベル認識部484は、修飾語検出部480から出力される文字列（修飾語に対応する文字

列）と要求レベル認識用データ格納部482に格納されているデータテーブルの内容に基づいて、利用者の要求レベルを認識し、認識結果判定部16から与えられる指示に基づいてこの認識結果を出力する。例えば、修飾語検出部480から出力された文字列が「スゴク（すごく）」であれば、要求レベル認識部484は、要求レベル認識用データ格納部482に格納されたデータテーブルに基づき、利用者の要求レベルがレベル3であると認識する。また、修飾語検出部480から所定の修飾語が検出されなかった旨の通知が出力された場合には、要求レベル認識部484は、利用者の要求レベルがレベル2であると認識する。

【0040】なお、上述した例では、利用者の要求レベルを判定するための要素として、声の大きさ（音圧レベル）、声の高さ、発声間隔、および所定の修飾語のいずれかの要素を考慮する場合について説明したが、これらの要素を2つ以上組み合わせることで利用者の要求レベルを判定するようにしてもよい。

【0041】また、上述した実施形態では、利用者の要求レベルをレベル1～3の3段階に分けていたが、この要求レベルは3段階に限定されるものではなく、簡単には2段階に設定してもよいし、より多段階に設定してもよい。

【0042】また、上述した実施形態では、本発明を適用した音声処理装置を車載用システムに適用した場合の一例について説明したが、本発明はこれに限定されるものではなく、種々のシステムに対して適用することが可能である。

【0043】

【発明の効果】上述したように、本発明によれば、操作対象と要求レベルが具体的に示されていないような音声が入力された場合でも、操作文字列の内容とこの操作文字列に関する要求レベルとに基づいて操作対象を特定して動作指示を行うことができるので、人間同士の会話に近い形の自然な言葉を用いて、各種の操作指示を入力することができるとともに、要求レベルに応じた適切な動作指示を選択することができる。

【図面の簡単な説明】

【図1】一実施形態の車載用システムの構成を示す図である。

【図2】欲求型文字列と具体指示型文字列の一例を示す図である。

【図3】要求レベル判定部の詳細構成を示す図である。

【図4】動作指示テーブルの一例を示す図である。

【図5】音声処理装置の動作手順を示す流れ図である。

【図6】声の高さの変化に基づいて利用者の要求レベルを判定する場合の要求レベル判定部の詳細構成を示す図である。

【図7】発声間隔の変化に基づいて利用者の要求レベルを判定する場合の要求レベル判定部の詳細構成を示す図

である。

【図8】変形例における音声処理装置の構成を示す図である。

【図9】音声認識部から出力される文字列に基づいて利用者の要求レベルを判定する場合の要求レベル判定部の詳細構成を示す図である。

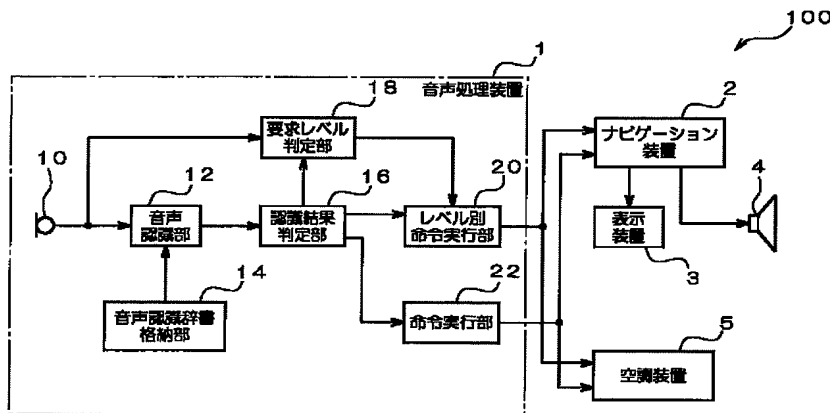
【図10】修飾語と利用者の要求レベルとの関係を示すデータテーブルの一例を示す図である。

【符号の説明】

- 1 音声処理装置
2 ナビゲーション装置

- 3 表示装置
4 スピーカ
5 空調装置
10 マイクロホン
12 音声認識部
14 音声認識辞書格納部
16 認識結果判定部
18 要求レベル判定部
20 レベル別命令実行部
22 命令実行部
100 車載用システム

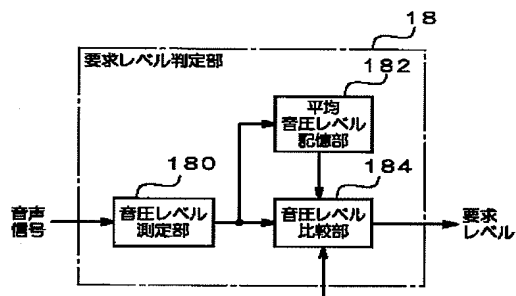
【図1】



【図2】

欲求型文字列	具体指示型文字列
ツカレタ (疲れた)	エアコンラッセル (エアコンをつける)
オナカガスイタ (お腹がすいた)	レストランラサガス (レストランを探す)
アツイ (暑い)	又ケミチラサガス (抜け道を探す)
サムイ (寒い)	
⋮	⋮
⋮	⋮

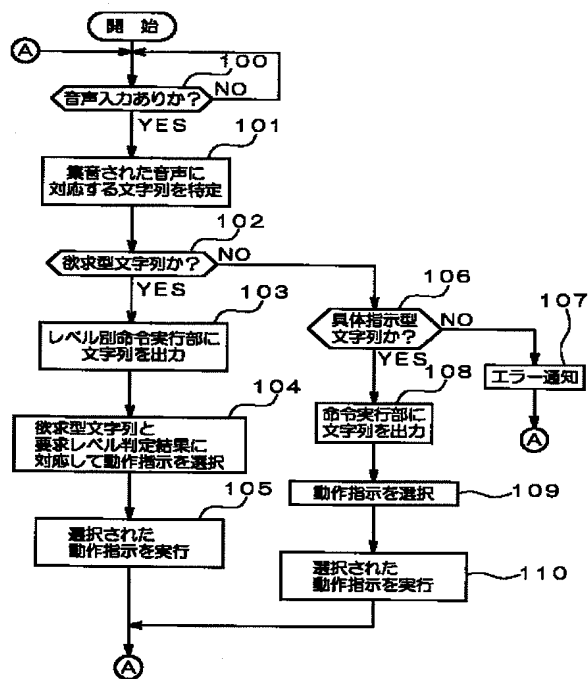
【図3】



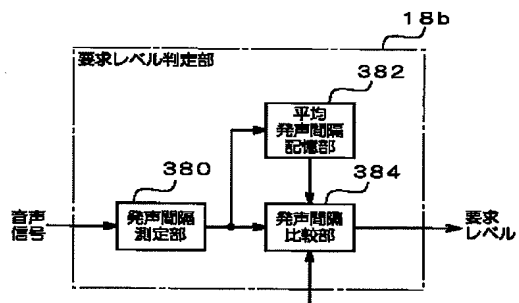
【図4】

欲求型文字列	レベル	動作指示命令
ツカレタ	1	連続運転時間の提示
	2	休憩を促すメッセージの提示
	3	休憩地の探索指示
アツイ	1	車室内温度の提示
	2	空調装置を動作させるかを問うメッセージの提示
	3	空調装置の動作指示
⋮	⋮	⋮
⋮	⋮	⋮

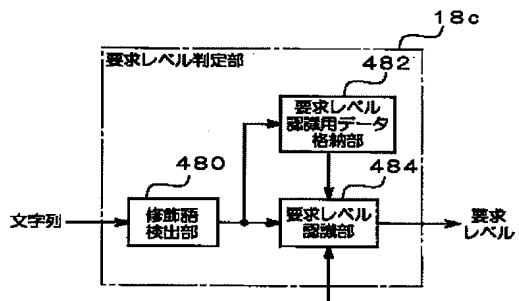
【図5】



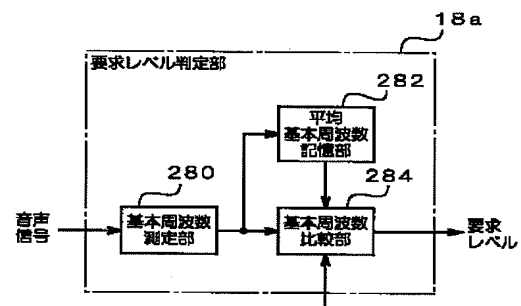
【図7】



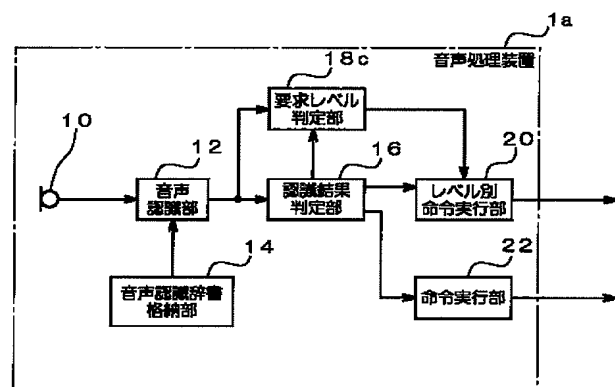
【図9】



【図6】



【図8】



【図10】

	レベル		
	1	2	3
修飾語	少し	X	すごく
	ちょっと		とても
	少々		かなり

フロントページの続き

(51)Int. Cl. 7	識別記号	F I	テーマコード (参考)
G 1 0 L 15/02		G 1 0 L 3/00	5 5 1 J
// G 1 0 L 101:00			5 5 1 Q
101:065			5 6 1 H
		9/00	3 0 1 A

F ターム (参考) 2F029 AA02 AC14 AC18
5D015 AA05 HH23 JJ00 KK01 LL10
5H180 AA01 FF22 FF25 FF27 FF32
9A001 HH17 JJ77 KK46